



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

THE ANALYST.

VOL. VI.

MAY, 1879.

No. 3.

ON THE DEVELOPMENT OF $[p + (1-p)]^n$.

BY E. L. DE FOREST, M. A., WATERTOWN, CONN.

It is well known that when p and q denote respectively the probabilities that an event will happen and that it will fail to happen, in a single trial, so that $p + q = 1$, then if m trials are to be made independently of each other, the $m+1$ terms of the development

$$(p+q)^m = p^m + \frac{m}{1} p^{m-1} q + \frac{m(m-1)}{1.2} p^{m-2} q^2 + \dots + q^m, \quad (1)$$

are the several probabilities that the event will happen any number of times from m down to 0, and that it will fail to happen a corresponding number of times from 0 up to m . (See for instance Merriman's *Least Squares*, p. 123.) The terms of the series increase in value from either extremity toward an intermediate point where they attain a maximum. If we regard them as equidistant ordinates to a curve, then as m is increased, the form of this curve approximates rapidly to the probability curve

$$y = ce^{-h^2 x^2},$$

and this is found to be the limiting form, when m becomes infinite. I propose here to demonstrate it in a manner rather more simple than any I have met with, and also to consider some points not usually noticed in connection with it.

Let the terms of the development be denoted by y_0, y_1, y_2 &c., so that

$$(p+q)^m = y_0 + y_1 + y_2 + \dots + y_m.$$

Comparing this with (1) we see that

$$\frac{y_1 - y_0}{y_0} = \frac{qm}{p} - 1, \quad \frac{y_2 - y_1}{y_1} = \frac{q(m-1)}{2p} - 1, \quad \frac{y_3 - y_2}{y_2} = \frac{q(m-2)}{3p} - 1,$$

and in general
$$\frac{y_{i+1} - y_i}{y_i} = \frac{q(m-i)}{p(i+1)} - 1. \quad (2)$$

When y , is denoted simply by y , and the constant interval between two consecutive ordinates is Δx , and the origin is taken at the first term, then

and (2) reduces to $y_{i+1} - y_i = \Delta y$, $x = i\Delta x$,

$$\frac{\Delta y}{y} = \frac{(qm-p)\Delta x - x}{p(x+\Delta x)}. \quad (3)$$

The largest term in the series (1) is that term whose rank is $i = qm$, as can be ascertained by trial in any particular case where qm is a whole number, and also by considering that the term in question is then

$$\frac{\Gamma(m+1)}{\Gamma(pm+1)\Gamma(qm+1)} p^{pm} q^{qm},$$

being the probability that the event will happen pm times and fail qm times, while the terms next preceding and following are

$$\frac{\Gamma(m+1)}{\Gamma(pm+2)\Gamma(qm)} p^{pm+1} q^{qm-1}, \quad \frac{\Gamma(m+1)}{\Gamma(pm)\Gamma(qm+2)} p^{pm-1} q^{qm+1}.$$

Dividing each of these two by the former, we find the quotients

$$\frac{pm}{pm+1} \text{ and } \frac{qm}{qm+1},$$

both of which are less than unity. We can always make qm a whole number by taking m sufficiently large.

In (3) let the origin be transferred from the first term in the series to the greatest term, by putting $x + qm\Delta x$ in place of x ; then

$$\frac{\Delta y}{y} = \frac{-(x+p\Delta x)}{p(qm\Delta x + x + \Delta x)}. \quad (4)$$

If m becomes infinite, and the successive ordinates are set close together so as to be consecutive, then at the limit Δx and Δy become dx and dy , and dx and pdx vanish in comparison with x , so that (4) reduces to

$$\frac{dy}{y} = \frac{-x}{p(qmdx + x)}. \quad (5)$$

To find the equation of the curve from this by integration, we must observe that the rules of the integral calculus presuppose that a differential is infinitely smaller than the function from which it is derived. If the numerator dy in the first member is to be infinitely smaller than the denominator y , it will follow that in the second member, the numerator x must be infinitely less than $qmdx + x$, and consequently the x in the denominator is infinitely less than $qmdx$, and may be neglected, so that we may write

$$\frac{dy}{y} = \frac{-x}{pqmdx}. \quad (6)$$

But since x is finite, it cannot be infinitely less than $pqmdx$ unless mdx is infinite, and this cannot be unless m is an infinity of the second order, that is, of a magnitude comparable to $1 \div (dx)^2$. To make it such, we must suppose that the series extends throughout the whole infinite length of the axis of X . For dx is the interval between any two consecutive terms or ordinates, and it is contained ∞ times in a finite distance, say the unit of x , and this unit is contained ∞ times in the whole axis of X , so that dx is contained ∞^2 times in the whole axis, and thus the whole number of terms in the series becomes $m+1 = \infty^2+1$. Le (6) then be written

$$\frac{dy}{y} = \frac{-xdx}{pqm(dx)^2} \quad (7)$$

The denominator of the second member does not vary with x , and $m(dx)^2$ is a constant and finite area. We will express the denominator by means of a new constant h , so that

$$h = 1 \div [dx \sqrt{(2pqm)}], \quad (8)$$

and (7) becomes

$$\frac{dy}{y} = -2h^2 x dx,$$

and integration gives $\log' y = -h^2 x^2 + \log' c$; $\therefore y = ce^{-h^2 x^2}$ (9)

Since $p+q = 1$, we have $(p+q)^m = 1$, and the sum of all the terms in the series (1) is unity, wherefore the sum of all the consecutive values of y in the curve (9) is also unity, if that curve fully represents the limit of the series. Multiplying the sum of the ordinates, and unity, each by dx , we have the condition

$$\int_{-\infty}^{\infty} y dx = dx; \therefore c \int_{-\infty}^{\infty} e^{-h^2 x^2} dx = dx. \quad (10)$$

The known value of the definite integral here is $\sqrt{\pi} \div h$. Thus (10) determines the constant

$$c = h dx \div \sqrt{\pi}, \quad (11)$$

and substituting it in (9), we get

$$y = \frac{h dx}{\sqrt{\pi}} e^{-h^2 x^2}, \quad (12)$$

the final equation of the limiting curve. We infer from (8) that h is the reciprocal of a finite line, so that $h^2 x^2$ is an abstract number, and finite as long as x is finite. It follows that y is an abstract number, as a probability should be, and is an infinitesimal of the first order.

Let i now be reckoned as x is, from the place of the maximum probability. As before, $x = i dx$ and at the limit $x = idx$. Substituting in (12) this value of x and also the value of h from (8), we find

$$y = (2\pi pqm)^{-\frac{1}{2}} e^{-i^2 \div 2pqm} \quad (13)$$

When an infinite number m of trials are made, the most probable result, as we have seen, is that the event will happen pm times and fail qm times, and (13) gives the probability that it will happen exactly $pm+i$ times and fail $qm-i$ times. This formula holds good approximately also when m is a large finite number, provided that i is considerably smaller than pm or qm , a proviso which results from the fact that m in the formula is an infinity of the second order, while i is an infinity of the first order, since $i = x \div dx$, and x is finite.

Since y is the probability that the number of times the event happens will deviate from the most probable number pm by a certain number i , we may regard $x = idx$ as a deviation of position, or error of position, and regard y in general as the probability that any observed quantity, subject to accidental causes of error, will deviate from its most probable value by a certain quantity or error x . Equation (12) then is the equation of the probability curve, expressing the law of facility of error.

Let both its members be divided by dx , which is equivalent to multiplying them by a constant and infinitely large quantity. Denoting $y \div dx$ by Y , we have

$$Y = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}, \quad (14)$$

and since y is an infinitesimal of the first order, Y is finite. This is what is commonly known as the equation of the probability curve, but it is evident that here the probability of the error x is not the ordinate Y , but the elementary area $Ydx = y$. The whole area of the curve is

$$\int_{-\infty}^{\infty} Ydx = 1.$$

The probability that any error which occurs, when taken without regard to sign, will not exceed a certain limit x , is the sum of the probabilities of all the errors from $-x$ up to $+x$, and equals the area of the curve between those limits, or

$$P = \frac{h}{\sqrt{\pi}} \int_{-x}^x e^{-h^2 x^2} dx. \quad (15)$$

If we put $hx = t$, then $dx = dt \div h$, the limits $x = \pm x$ become $t = \pm hx$, and Y being a function of x^2 the curve is symmetrical on each side of the origin, so that

$$P = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt. \quad (16)$$

All works on probability contain the values of this integral arranged in a table, enabling us to find P for any given value of the argument $t = hx$.

It follows from (8) also that

$$t = i \div \sqrt{(2pqm)}, \quad (17)$$

so that such a table will give the sum of $2i$ values of y in (13), being the approximate probability that when a large finite number m of trials are made, and i is an integer or half-integer much smaller than pm or qm , the number of times the event will happen will be one of the group of $2i$ consecutive integers the middle of which falls nearest to the number pm .

We have hitherto supposed the constant h to be determined from (8) in terms of m , but in applying the theory of probability of errors it generally happens that there are given only a certain finite number of observed errors or deviations from the true value of a quantity or from some value which is regarded as the most probable one, while the whole number m of different possible errors is infinite or at least unknown. To find h from the observed errors, the best course generally is to compute the "mean error", or as it has been more properly called, the *quadratic error*, which is defined to be the square root of the mean of the squares of all the observed errors. In the most typical case, when the number of observed errors is very large or infinite, the square of the quadratic error is the quotient arising from dividing the sum of the squares of all the errors by the whole number of errors, that is

$$\epsilon^2 = \int_{-\infty}^{\infty} x^2 Y dx \div \int_{-\infty}^{\infty} Y dx, \quad (18)$$

each error x being supposed to occur a number of times proportional to its probability $Y dx$. The value of the divisor here is unity, as we have seen. The dividend is

$$\frac{h}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^2 e^{-h^2 x^2} dx = \frac{1}{h^2 \sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt.$$

This last definite integral is known to have the value $\frac{1}{2}\sqrt{\pi}$. The quadratic error then is

$$\epsilon = 1 \div h \sqrt{2}; \quad \therefore h = 1 \div \epsilon \sqrt{2}. \quad (19)$$

Since ϵ^2 is the mean of all the observed values of x^2 , and $x = ix$, it follows that $(\epsilon \div dx)^2$ is the mean of all the observed values of i^2 . Substituting in (19) the value of h from (8), we get

$$\epsilon = dx \sqrt{(pqm)}; \quad \therefore \epsilon \div dx = \sqrt{(pqm)}. \quad (20)$$

This is the square root of the mean of the values of i^2 in (13), bearing the same relation to deviations i from the most probable number pm , as ϵ bears to deviations x from the most probable value of any observed quantity. It is of use in statistical investigations. (*Smithsonian Report* of 1873, p. 334.)

The constant h has been called the "measure of precision", because it varies inversely as the quadratic error; but in point of fact the accuracy of a

set of observations is usually estimated by the *probable error*, defined to be that error r , which occupies an intermediate position among all the observed errors taken without regard to sign, so that for an infinite number of errors the number greater than r is equal to the number less than r , and the probability that any error which occurs will be, for example, less than r , is $\frac{1}{2}$. By interpolation in a table of values of P , the value of $t = hx$ corresponding to $P = \frac{1}{2}$ is found to be

$$hr = 0.4769,$$

and substituting the value of h from (19) we get

$$r = .6745\epsilon, \quad (21)$$

by which the probable error is computed from the quadratic error. We also have by (20)

$$r \div dx = .6745\sqrt{(pqm)}, \quad (22)$$

which is the probable or intermediate value of i , so that it is an even chance that the event will happen a number of times not deviating from the most probable number pm by more than this probable value, when a large number m of trials are made.

It should be noticed, however, that the series is in general symmetrical on both sides of the maximum probability or axis of Y , only when m is infinite. If m is finite, the terms of the series are not identical on opposite sides of that axis, and the probable error or deviation is somewhat greater on one side, and less on the other, than the value which (22) gives. It is greatest on that side which is toward the more remote end of the series. But in the special case where $p = q = \frac{1}{2}$, the maximum term is in the middle of the series, and the probable errors are alike on both sides, whether m be infinite or finite. One of the usual elementary proofs of the law of probability demonstrates it as the limit of the development of $(\frac{1}{2} + \frac{1}{2})^{2n}$. (See for instance Price's *Calculus*, Vol. II., p. 376.)

Let us now return and consider the validity of our treatment of formula (5), from which (6) was obtained by neglecting the finite x in comparison with the infinite $qmdx$, where mdx represents the whole length of the axis of X , or the whole distance occupied by the series (1) when m is an infinity of the second order. It is evident, in the first place, that dy is not infinitely smaller than y in every part of the series. On the contrary, it is sometimes of a magnitude equal or comparable to that of y , and sometimes even infinitely greater than y , as appears from (5).

When we put successively $x = -qmdx + a$, $x = pmdx - a$, where a is finite and represents distance from either end of the series, (5) becomes

$$\frac{dy}{y} = \frac{qmdx - a}{ap} \quad \text{and} \quad \frac{dy}{y} = -\frac{(pmdx - a)}{(pmdx - ap)}.$$

As mdx is infinite, $dy \div y$ is infinite in the one case and -1 in the other, showing that, of two consecutive ordinates at any finite distance from either end of the series, that one which is nearest the end is infinitely smaller than the other one. Therefore, in going from either end toward the middle, the ratio of dy to y passes from infinity at any finite distance from the end, through intermediate and finite values, down to infinitesimal values for any finite distance from the point where y is a maximum; and it is only to this latter region that formula (6) applies. Have we a right, then, to say that the terms outside of this region are so small that their sum is infinitesimal, and that the sum of the terms within it is therefore equal to unity? For we assumed this, when we determined the constant c in (9) by the condition (10). It would hardly seem to be self-evident, but there are facts which confirm its truth, and one of them is reached by computing the quadratic error for a finite series, as follows.

In (1) let p , q and m have any consistent values, for instance

$$p = \frac{1}{8}, \quad q = \frac{2}{8}, \quad m = 4,$$

so that

$$(p+q)^m = \frac{1}{81}(1+8+24+32+16).$$

The second member is a series of five fractions which are the probabilities of the five different possible results, and may be represented by equidistant ordinates, the constant interval between them being Δx . Let the origin be taken at the point whose distance from the first ordinate is $qm\Delta x = \frac{8}{3}\Delta x$. This point falls two thirds of the way from the term 24 to the next term 32. The negative errors or deviations then are

$$-\frac{2}{3}\Delta x, \quad -\frac{5}{8}\Delta x, \quad -\frac{8}{8}\Delta x,$$

and the positive ones are $\frac{1}{8}\Delta x, \quad \frac{4}{8}\Delta x$.

The square of the quadratic error is the mean of the squares of all the errors, each one being supposed to occur a number of times proportional to y , so that

$$\epsilon^2 = \frac{\frac{1}{81}[1(\frac{8}{3})^2 + 8(\frac{5}{8})^2 + 24(\frac{2}{8})^2 + 32(\frac{1}{8})^2 + 16(\frac{4}{8})^2](\Delta x)^2}{\frac{1}{81}(1+8+24+32+16)},$$

from which we find

$$\epsilon^2 = \frac{8}{81}(\Delta x)^2 = pqm(\Delta x)^2.$$

In like manner, for all other consistent values of p , q and m which have been tried, it has been found that the quadratic error is invariably $\Delta x \sqrt{pqm}$. Hence we infer that at the limit, when m is an infinity of the second order, and Δx is reduced to dx

$$\epsilon = dx \sqrt{pqm}.$$

But this is the value which has already been found in (20) by integrating from the probability curve. In other words, the summation of x^2y for the whole of the infinite series gives the same result as its summation from the

curve (12) which can represent truly only a portion of the series; while the sum of y is unity in both cases. We conclude then that in all those portions of the infinite series to which (6) and (12) do not apply, the terms y are so small that the total of x^2y , and consequently the total of y , is infinitesimal.

This method of finding the expression (20) for the quadratic error inductively from the finite series is one which I have not seen in any of the published works on probability.

I take this opportunity to make a correction in my article on the Limit of Repeated Adjustments (ANALYST, Sept. 1878), where the curve, whose ordinates y are the coefficients of the limiting resultant formula of adjustment, was shown to be a probability curve, or a higher curve of analogous nature, and the condition that the sum of all the consecutive ordinates must be unity was employed in determining the constant parameters of the curves.

This condition was regarded as being expressed by

$$\int_{-\infty}^{\infty} y dx = 1,$$

whereas it should have been

$$\frac{1}{dx} \int_{-\infty}^{\infty} y dx = 1,$$

as we have seen in connection with formula (10) of the present paper. The consequence is, that the expressions for y obtained in formulas (10), (14) and (19) of the article referred to, are really expressions for $Y = y \div dx$, and the numerical values given for y in Tables I. and III. are really values of Y . The correction is one that affects the theory of the matter rather than the practical results obtained. It has been a fault of long standing with writers on the probability curve, to slur over the distinction between the infinitesimal probability here denoted by $y = Ydx$, and the finite ordinate Y which is proportional to y . When y is regarded as an ordinate, it belongs to a curve whose total area is dx ; but Y belongs to one whose area is unity.

It may be well to add a remark upon the three formulas (6), (7) and (8) of the article referred to, where k represents the number of applications of the adjusting process, and is supposed to become infinite at the limit. The infinity in these three cases must be of the second, fourth and sixth orders respectively, as appears from considerations similar to those arising in formula (6) of the present paper, where m becomes an infinity of the second order. There cannot well be any limit to the increase of an abstract number. If the number ∞ is a supposable quantity, then ∞^2 , ∞^4 and ∞^6 are supposable also.

For example, if dx is contained ∞ times in the unit of x , and this unit of x is contained ∞ times in the whole axis of X , then dx is contained ∞^2 times in the whole axis of X , and $(dx)^2$ is contained ∞^4 times in the infinite plane XY , and $(dx)^3$ is contained ∞^6 times in the infinite space XYZ .

The concrete result in the three cases of adjustment appears to be this, that for a finite number of applications, the limiting form is approached more rapidly in the first case than in the second, and more rapidly in the second than in the third. Thus under the first case, p. 133, the coefficients for only 16 applications approach the limiting form more nearly than the coefficients for 32 applications do under the second case at p. 138, the limit being apparently farther off in the latter case.

NOTE.—When a series is adjusted by two formulas successively, as at p. 66 of ANALYST for May 1878, where the two are denoted by

$$u'_0 = l_0 u_0 + l_1(u_1 + u_{-1}) + l_2(u_2 + u_{-2}) + \dots + l_m(u_m + u_{-m}),$$

$$u''_0 = L_0 u'_0 + L_1(u'_1 + u'_{-1}) + L_2(u'_2 + u'_{-2}) + \dots + L_n(u'_n + u'_{-n}),$$

the coefficients in the resultant formula are the same as those which belong to the several powers of x in the product of the two polynomials

$$l_m + l_{m-1}x + l_{m-2}x^2 + \dots + l_0x^m + l_1x^{m+1} + \dots + l_mx^{2m},$$

$$L_n + L_{n-1}x + L_{n-2}x^2 + \dots + L_0x^n + L_1x^{n+1} + \dots + L_nx^{2n}.$$

This property, which I did not perceive at first, systematizes the matter and simplifies it to the mind, although it makes no difference in the actual amount of computation to be done, in forming the resultant coefficients.

Since the resultant for k successive adjustments by one formula has the same coefficients as the corresponding polynomial raised to the k power, we are enabled to demonstrate the proposition (3) at p. 129 of the ANALYST for Sept., 1878, by means of the Multinomial Theorem, as I will take occasion to show hereafter. The demonstration includes the general case where the adjustment formula is not symmetrical on each side of the middle, so that λ_+ and λ_- are not necessarily equal.

GENERAL RULES FOR THE FORMATION OF MAGIC SQUARES OF ALL ORDERS.

BY PROF. A. B. NELSON, CENTRE COLLEGE, DANVILLE, KY.

THE following general rules for the formation of magic squares, whether of odd or even degree, may be new to the younger generation of mathematicians, and possibly to some others. The squares constructed by these